

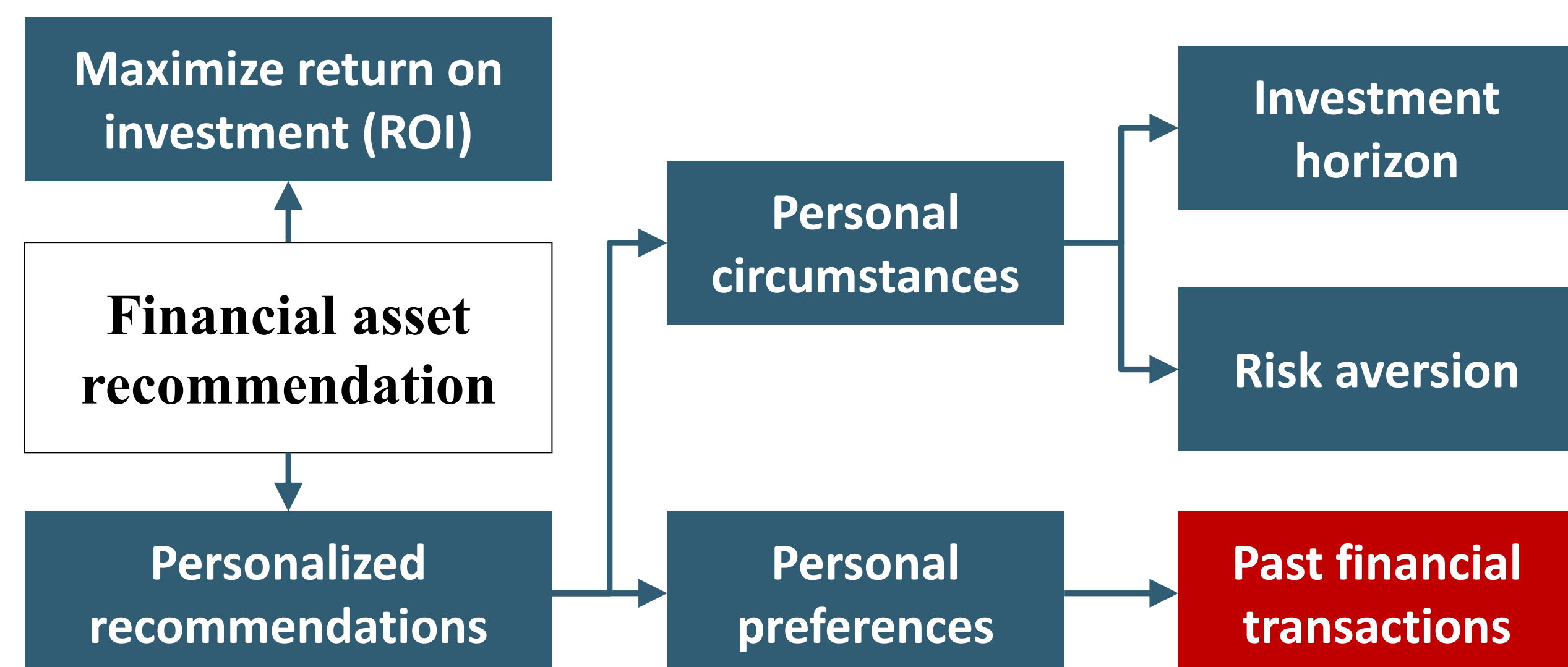
Overview

Financial Asset Recommender (FAR) systems use models trained on past data to suggest investment assets to customers. The effectiveness of FAR models can be evaluated in two ways: (a) measuring the money customers could obtain if they followed the recommendations (profitability-based) and (b) quantifying the ability of models to predict future customer investments (transaction-based). In this work we compare these strategies to determine which should be the primary metric for evaluating FAR systems.



Paper

1. Task and Motivation



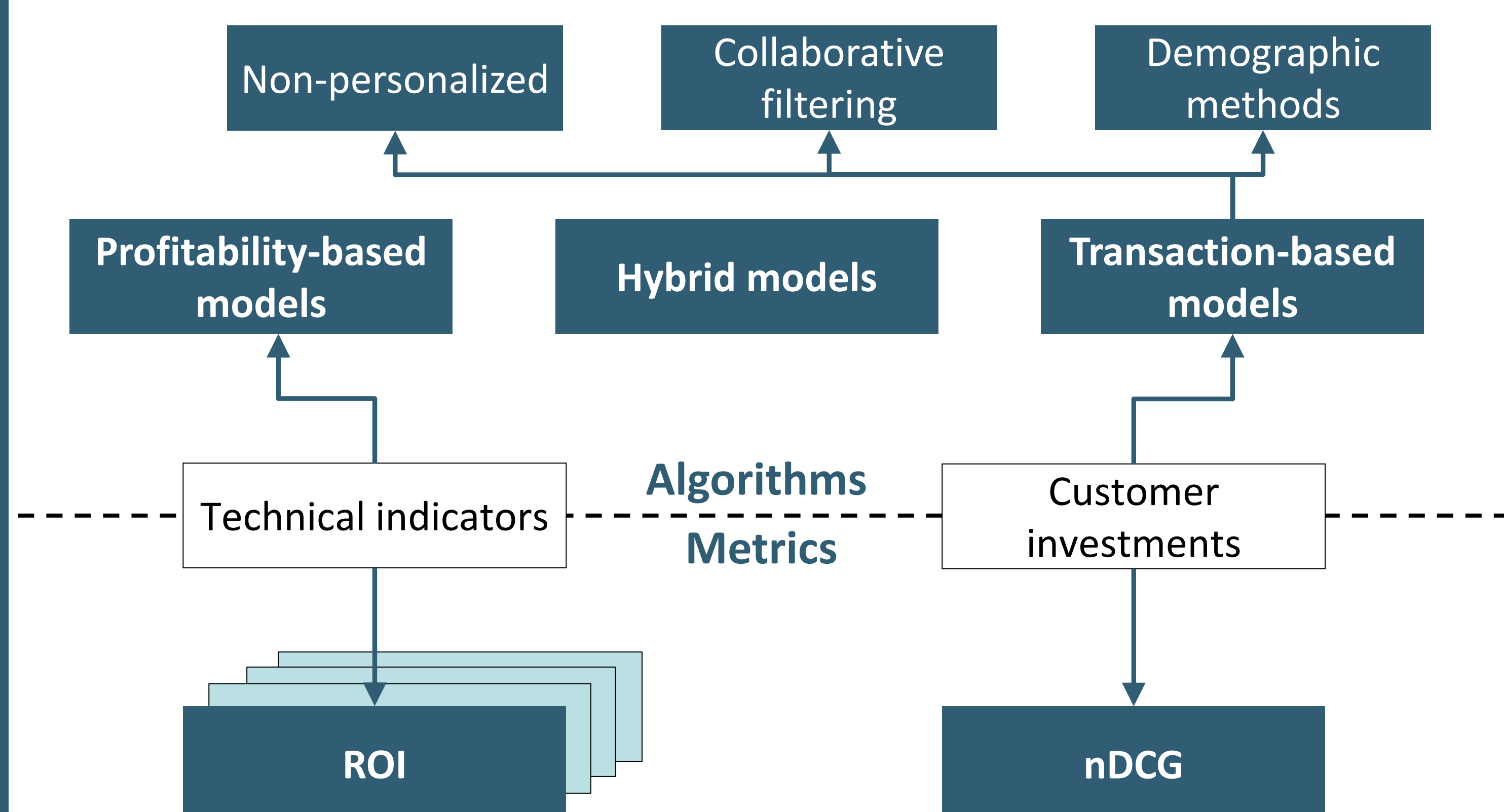
Transaction-based metrics

- calculate similarity between a set of recommendations and what the customer actually invested in (financial transactions) as a quality metric over the produced recommendations
- would be expected to be positively correlated with return on investment if customers are effective investors.
- would be a better aggregate measure of a FAR system, as they are inherently personalized if this is true

Research Question: Are transaction-based metrics to be positively correlated with return on investment and if not, why not?

2. Experiment Design

- We train a wide range of FAR systems that consider different types of input data to evaluate



- We compare how these models perform if the customer invests for 6 months under both transaction and profitability metrics

3. Dataset

- Snapshot of the Greek market covering a range of different securities: stocks, bonds, mutual funds and other banking products for the period between January 2018 and March 2021
 - Evaluation performed for 29 time points, spaced 2 weeks apart
 - Starting from 1st July 2019
 - Data previous to each time point as training, six months after as test

Market data		Customer data	
Property	Value	Property	Value
Unique assets	5,371	Unique customers	52,390
Assets with investments	2,025	Transactions	313,004
Price data points	1,768,128	Acquisitions	269,031
Average return (by assets, whole period)	23.67%	Average return (by customers, whole period)	18.41%
% profitable assets	53.08%	% customers with profits	58.00%

4. Results

- RQ1:** Are transaction-based and profitability-based metrics interchangeable when evaluating financial asset recommendation systems?

		Transaction-Based	Profitability-based	
Data source	Category	Algorithm	nDCG@10	ROI@10
None		Random	0.0223	0.0118
Technical indicators	Regression	SVR	0.0041	0.1212
		LightGBM	0.0599	0.1423
		Random forest	0.0570	0.0583
Past investments	Non-personalized	Popularity	0.3374	-0.0628
		LightGCN	0.3081	-0.0643
	Collaborative filtering	ARM	0.2687	-0.0647
		MF	0.0812	-0.0460
		UB kNN	0.1428	-0.0344
	Demographic	CPS	0.3003	-0.0544
Hybrid		Hybrid-nDCG	0.2454	-0.0466
		Hybrid-regression	0.0220	0.0382
Market average			0.1026	

Best ROI / Worst nDCG

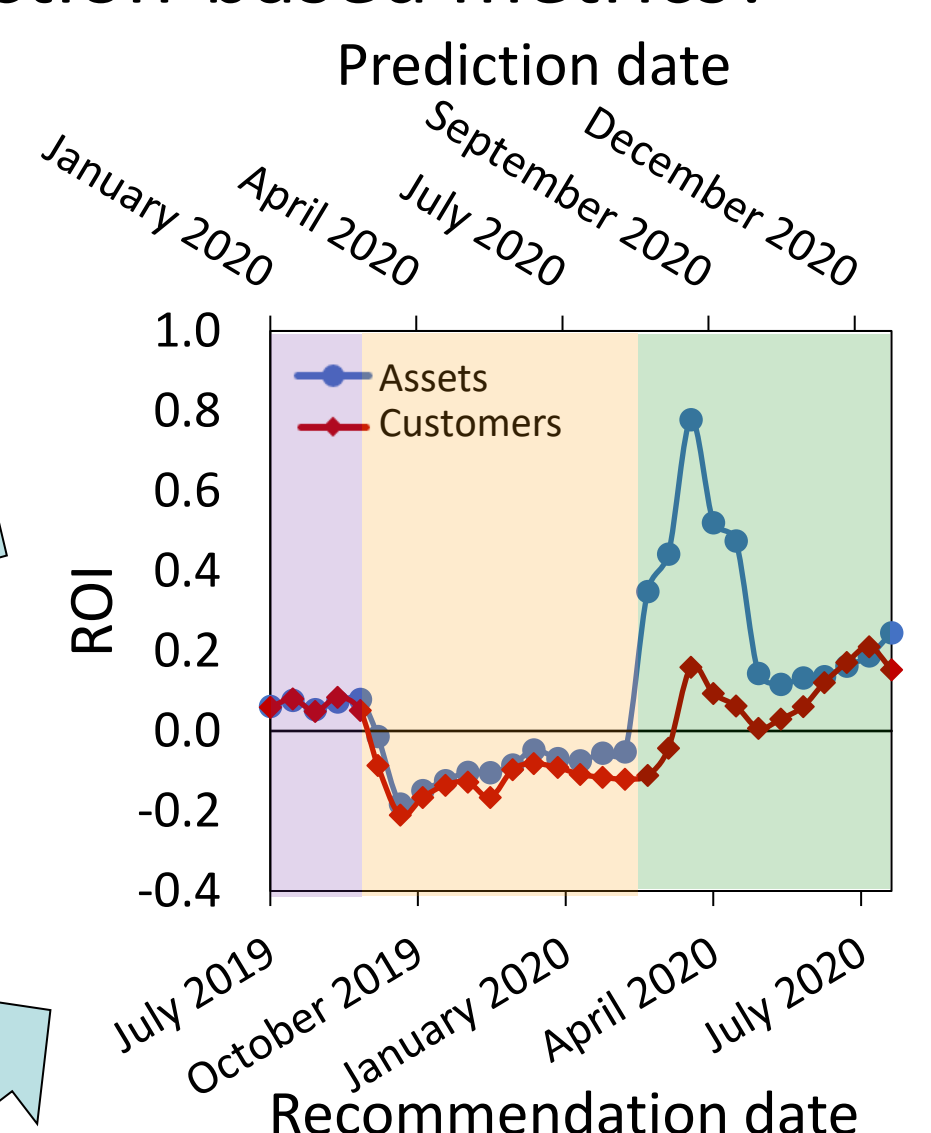
Best nDCG / Worst ROI

- No**, profitability-based metric and transaction-based metrics are not correlated – models can perform well under transaction-based metrics but still lose money

- RQ2:** What are the main factors that influence transaction-based metrics?

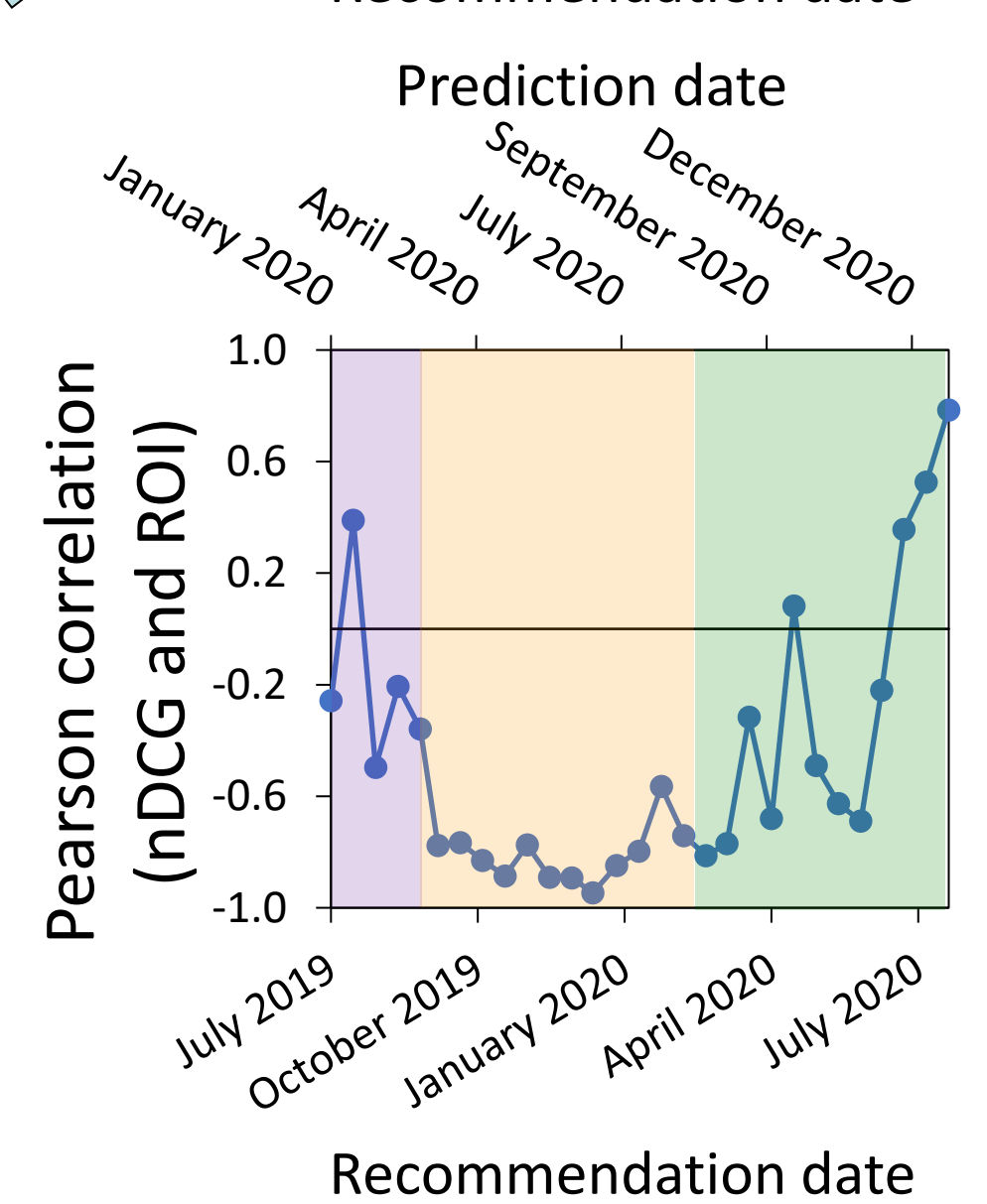
- Are our customers poor investors?**

- On average the customers in this dataset underperform the market; this seems to be exacerbated by high volatility in the market



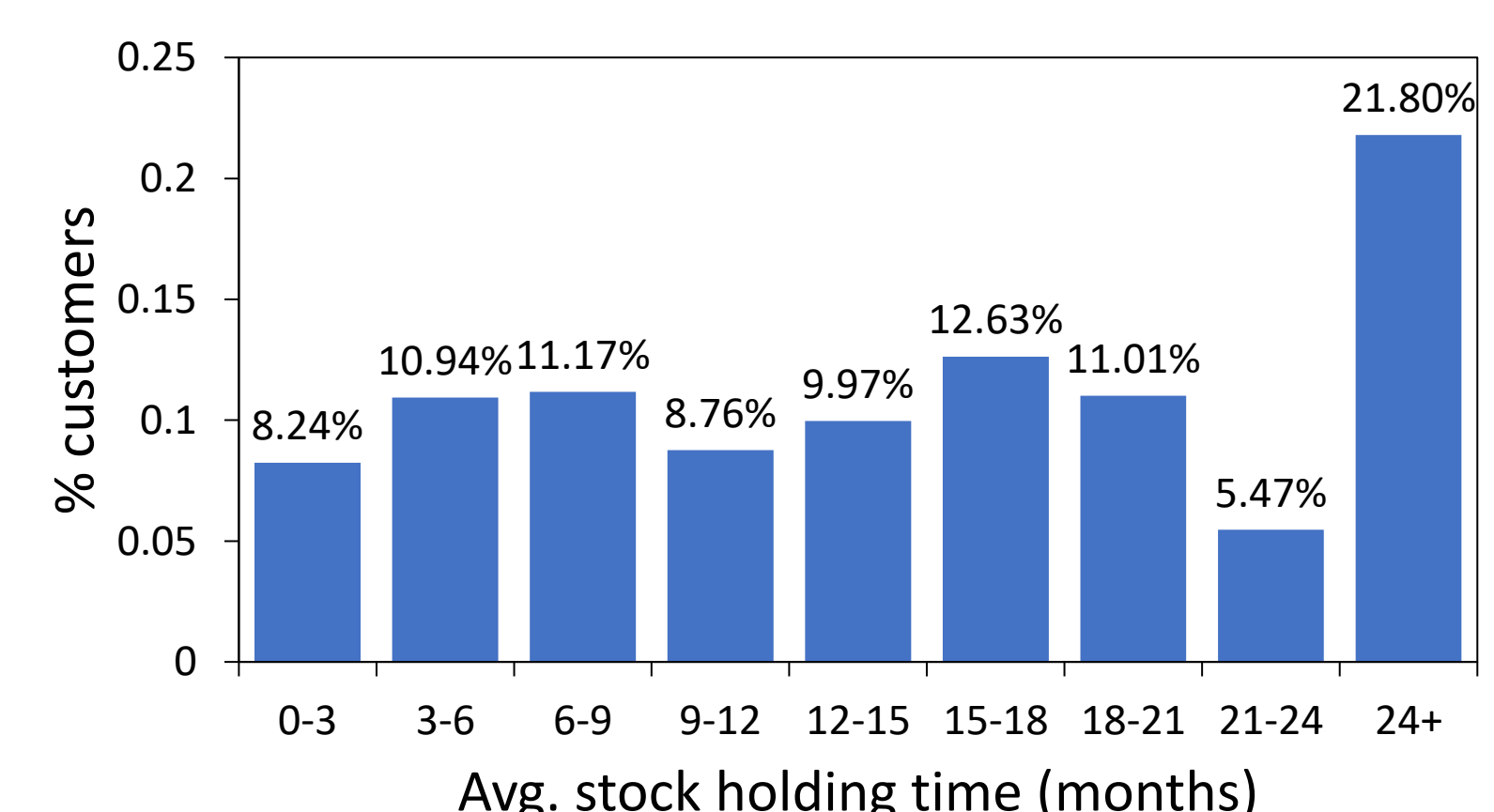
- Was the time period of this dataset a-typical?**

- We can see this unusual fall in profitability in late 2019 (COVID-19 impact, in yellow)
- Also, a spike in profitability when predicting for early 2020 (COVID-19 recovery, in green)
- There were far fewer profitable assets during the downturn, followed by many profitable assets during the market rally
- During these downturn, volatile times, transactions and profitability become negatively correlated



- Is ROI after 6 months a bad profitability metric?**

- Only for a subset of the customers, we can see that there is a wide spread in terms of holding time amongst the customers, indicating that a range of prediction time horizons are needed



5. Conclusion

Multiple factors such as variable investment horizons, market volatility and the ability of the customers to navigate the market make transaction-based metrics risky as an evaluation metric for FAR systems.



This work was carried out as part of the Infinitech project which is supported by the European Union's Horizon 2020 Research and Innovation programme under grant agreement no. 856632.